On optimal entanglement assisted one-shot classical communication

Brett Hemenway,[1,*] Carl A. Miller,[2,†] Yaoyun Shi,[2,‡] and Mary Wootters[1,§]

[1] *Mathematics Department, University of Michigan, Ann Arbor, MI 48109, USA*
[2] *Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA*

(Dated: December 27, 2012)

The *one-shot success probability* of a noisy classical channel for transmitting one classical bit is the optimal probability with which the bit can be sent via a single use of the channel. Prevedel *et al.* (*PRL* **106**, 110505 (2011)) recently showed that for a specific channel, this quantity can be increased if the parties using the channel share an entangled quantum state. We completely characterize the optimal entanglement-assisted protocols in terms of the radius of a set of operators associated with the channel. This characterization can be used to construct optimal entanglement-assisted protocols from the given classical channel and to prove the limit of such protocols. As an example, we show that the Prevedel *et al.* protocol is optimal for two-qubit entanglement. We also prove some simple upper bounds on the improvement that can be obtained from quantum and no-signaling correlations.

Suppose that two parties, Alice and Bob, communicate over a noisy classical channel. While there are many examples of how Alice and Bob may benefit when they upgrade to a quantum channel, examples in which shared entanglement improves communication over a classical channel have only recently been discovered [1–3]. That these examples exist at all is somewhat surprising, as neither shared entanglement [4] nor the assistance of non-signaling correlations [3] can increase the classical capacity of the channel. So far, work in this direction has focused on the the *(one-shot) zero error capacity*, which measures the number of messages Alice can send to Bob perfectly [3, 5–7], and the related notion of the *one-shot success probability* [2], which is the best probability with which Alice can successfully send a single bit to Bob.

It is of interest to determine how shared entanglement affects these two quantities, as this will further our understanding of how resources from quantum mechanics can be used for communication.

Previous work on enhanced communication over a classical channel has focused on the assistance that can be provided by non-signaling correlations. In this setting, both the zero error capacity and one-shot success probability can be written as the solution to linear programs [3, 5]. Some upper bounds are known for the entanglement assisted zero error capacity [6]; these bounds are often the best bounds available in the unassisted case, suggesting that there are strong limitations to the amount of assistance that entanglement can provide. Much less is known about the limits of quantum assistance for the one-shot success probability. In [2], Prevedel *et al.* give an example of a channel where the unassisted success probability, $\mathrm{Succ}(N)$, the entanglement-assisted success probability $\mathrm{Succ}_{\mathrm{Q}}(N)$, and the non-signaling assisted success probability $\mathrm{Succ}_{\mathrm{NS}}(N)$ are all different. It is known that entanglement cannot be completely helpful: if $\mathrm{Succ}(N)$ is less than one, then so is $\mathrm{Succ}_{\mathrm{Q}}(N)$ [3]. However, the size of the gap between them has remained unquantified.

We use two distinct approaches to quantify the extent to which entanglement can help Alice and Bob. In our first approach, we derive a simple formula for $\mathrm{Succ}_{\mathrm{Q}}(N)$ in terms of the dimension of the entanglement. This formula, which is given by maximizing a quantity over a family of positive semidefinite operators, is easy to work with, and as an example of its applicability, we show that the protocol from [2] is in fact optimal for their channel and for 2-dimensional entanglement assistance.

While our first approach is quite general, it does not give a closed form for the success probability. Our second approach obtains explicit closed-form upper bounds for the success probability. As a first step, we prove the following general bound on non-signaling assistance. Let r be the number of elements in the input alphabet of N. Then,

$$\frac{\mathrm{Succ}_{\mathrm{NS}}(N) - \frac{1}{2}}{\mathrm{Succ}(N) - \frac{1}{2}} \leq 2 - \frac{2}{r}. \tag{1}$$

The quantity $(\mathrm{Succ}(N) - \frac{1}{2})$ measures the advantage that Alice and Bob have over a random strategy; thus, (1) measures the additional advantage gained by non-signaling correlations. Our proof of (1) uses the linear program characterization of $\mathrm{Succ}_{\mathrm{NS}}(N)$ from [5]. From this, we derive an upper bound on the amount of assistance from a binary quantum device; we use the fact that any quantum correlation can be decomposed into a probabilistic mixture of a local correlation and a non-signaling correlation (the concept of *local fraction*). We show that both of these bounds are the best possible, in the sense that there are channels for which equality is achieved.

A common thread in both approaches above is the use of the *radius* of a subset of a normed vector space. Our formula for $\mathrm{Succ}_{\mathrm{Q}}(N)$ depends on maximizing the radius of a family of Hermitian operators. In the second approach we use a formula for $\mathrm{Succ}_{\mathrm{NS}}(N)$ (an alternate formulation of Proposition 14 from [5]) which is expressed in terms of the radius of a particular set of vectors.

Notation and terminology. Throughout this paper, we assume that Alice is trying to transmit a single bit to Bob across a classical channel. Alice and Bob will have access to a two-part input output device D (Figure 1), which may be classical, quantum, or implement an arbitrary non-signaling correlation. Each two-part input

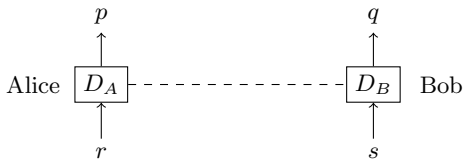

FIG. 1. A two-part input output device.

output device D gives rise to a *correlation* between Alice and Bob, given by

$$\{D(pq|rs) \mid p \in \mathcal{P}, q \in \mathcal{Q}, r \in \mathcal{R}, s \in \mathcal{S}\}$$

so that $D(pq|rs)$ is the probability of outputs p and q given inputs r and s. We will abuse notation by identifying the device D with the correlation it induces.

We say a device is *non-signaling* if the partial sums $\sum_{p\in\mathcal{P}} D(pq|rs)$ do not depend on r, and the partial sums $\sum_{q\in\mathcal{Q}} D(pq|rs)$ do not depend on s. We say a non-signaling device D is *quantum* if there exist Hilbert spaces V_A and V_B, families of POVMs $\{\{A_r^p\}_p\}_r$ $\{\{B_s^q\}_q\}_s$, and a density operator Λ on $V_A \otimes V_B$ such that $D(rs|pq) = \mathrm{Tr}((A_r^p \otimes B_s^q)\Lambda)$. The device is quantum with *dimension* n if both V_A and V_B are n-dimensional, and *binary* if the input and output alphabets have size 2.

A classical channel N is given by a matrix of conditional probabilities $\{N(y|x) \mid x \in \mathcal{X}, y \in \mathcal{Y}\}$, where $N(y|x)$ is the probability of seeing an output $y \in \mathcal{Y}$ given the input $x \in \mathcal{X}$. For any channel N, let $\mathrm{Succ}(N)$ denote the maximum probability with which a single bit can be sent across N (without assistance). Let $\mathrm{Succ}(N, D)$ denote the maximum probability for a single-bit transmission across N with the assistance of D. If $\mathbf{S}$ is a set of two-part devices, write $\mathrm{Succ}_{\mathbf{S}}(N) := \sup_{S\in\mathbf{S}} \mathrm{Succ}(N, S)$. We will be concerned with three choices of $\mathbf{S}$. We consider the set NS of non-signaling devices; the sets Q and $\mathrm{Q}(n)$ of quantum and n-dimensional quantum devices; and the set Q_b of binary quantum devices.

General quantum devices. In this section, we derive a formula for $\mathrm{Succ}_{\mathrm{Q}(n)}(N)$, and give an example of how to use our formula. We will use the *radius* of a finite set $\{H_i\}_{i\in\mathcal{I}}$ of Hermitian operators on a finite-dimensional Hilbert space V, defined by $\mathrm{Rad}\{H_i\}_i := \min_C \max_i \|H_i - C\|$ where the minimum is taken over all Hermitian operators C on V. The following lemma, which is proved in section 1 of the supplementary material, gives an alternative expression for the radius.

Lemma 1. *For any finite set $\{H_i\}_{i\in\mathcal{I}}$ of Hermitian operators on a finite-dimensional Hilbert space V, the radius of $\{H_i\}$ is equal to*

$$\max_{\substack{\lambda_i\geq 0, \lambda_i'\geq 0 \\ \sum\lambda_i = \sum\lambda_i' \\ \mathrm{Tr}(\sum\lambda_i)=1/2}} \left[\sum_{i\in\mathcal{I}} \mathrm{Tr}\left((\lambda_i - \lambda_i')H_i\right)\right].$$

Here, the maximization is over all Hermitian operators $\{\lambda_i\}_{i\in\mathcal{I}}$ and $\{\lambda_i'\}_{i\in\mathcal{I}}$ on V satisfying the given constraints.

Using this lemma, we will prove the following theorem, which characterizes $\mathrm{Succ}_{\mathrm{Q}(n)}(N)$.

Theorem 2. *For any channel N, and any integer $n \geq 2$,*

$$\mathrm{Succ}_{\mathrm{Q}(n)}(N) = \frac{1}{2} + \max_{\{B_y\}} \left(\mathrm{Rad}\left\{\sum_{y\in\mathcal{Y}} N(y|x)B_y\right\}_{x\in\mathcal{X}}\right),$$

where the maximization is over all families $\{B_y\}_{y\in\mathcal{Y}}$ of Hermitian operators on $\mathbb{C}^n$ satisfying $0 \leq B_y \leq \mathbb{I}$.

Proof. Consider the following quantum-assisted protocol for transmitting a single bit across N. Alice and Bob possess a bipartite quantum system represented by a density matrix Λ on a Hilbert space $V_A \otimes V_B$. Alice wishes to transmit a message $a \in \{0, 1\}$. Depending the value of a, she applies one of two possible POVMs $\{A_0^x\}_{x\in\mathcal{X}}$ or $\{A_1^x\}_{x\in\mathcal{X}}$ to V_A and sends the result of the measurement to the channel N. Bob receives the output y of the channel, and according to this output, applies one of a family of binary POVMs $\left\{\{B_y^0, B_y^1\}\right\}_{y\in\mathcal{Y}}$ to V_B. The result of this output is Bob's guess at Alice's original message.

In order to compute the success probability for this protocol, it is not necessary to know the state Λ or the operators $\{A_a^x\}_{x,a}$: it is only necessary to know the operators $\rho_a^x := \mathrm{Tr}_A\left[(A_a^x \otimes \mathbb{I})\Lambda\right]$, which represent the state of Bob's quantum system when the outcome of Alice's measurement is x. These operators satisfy $\sum_x \rho_0^x = \sum_x \rho_1^x$ and $\mathrm{Tr}(\rho_0^x) = 1$, and, in fact, any family of operators satisfying those two constraints can be induced by an appropriately chosen state Λ and appropriately chosen POVMs $\{A_0^x\}_{x\in\mathcal{X}}$ or $\{A_1^x\}_{x\in\mathcal{X}}$. Thus, for our purposes, to specify an (n, n)-dimensional entanglement-assisted strategy for communicating a single bit across N, it suffices to specify a collection of binary POVMs $\left\{\{B_y^0, B_y^1\}\right\}_{y\in\mathcal{Y}}$ on $\mathbb{C}^n$ and a collection of positive semidefinite operators $\{\rho_a^x \mid a \in \{0, 1\}, x \in \mathcal{X}\}$ on $\mathbb{C}^n$ satisfying

$$\sum_x \rho_0^x = \sum_x \rho_1^x \qquad \text{and} \qquad \mathrm{Tr}(\sum_x \rho_0^x) = 1. \tag{2}$$

The success probability of the protocol is given by

$$\frac{1}{2}\left(\sum_{y\in\mathcal{Y}} N(y|x) \sum_{x\in\mathcal{X}} \mathrm{Tr}(\rho_0^x B_y^0)\right) + \frac{1}{2}\left(\sum_{y\in\mathcal{Y}} N(y|x) \sum_{x\in\mathcal{X}} \mathrm{Tr}(\rho_1^x B_y^1)\right).$$

Let $B_y = B_y^0$. Since $B_y^1 = \mathbb{I} - B_y$, the expression above simplifies to

$$\frac{1}{2} + \frac{1}{2} \cdot \mathrm{Tr}\left[\sum_{x\in\mathcal{X}} (\rho_0^x - \rho_1^x) \sum_{y\in\mathcal{Y}} N(y|x) B_y\right]. \qquad (3)$$

The quantity $\mathrm{Succ}_{\mathrm{Q}(n)}(N)$ is the maximum of this expression over all $n \times n$ Hermitian operators $\{B_y\}_{y\in\mathcal{Y}}$ satisfying $0 \leq B_y \leq \mathbb{I}$ and all $n \times n$ positive semidefinite operators $\{\rho_a^x\}_{x\in\mathcal{X}, a\in\{0,1\}}$ satisfying (2) above. Applying Lemma 1 yields the desired formula. □

A convexity argument (see section 2 of the supplementary information) proves the following stronger version of Theorem 2.

Corollary 3. *The formula in Theorem 2 holds also when the maximum is taken only over families $\{B_y\}$ that consist of projections on $\mathbb{C}^n$.*

As an example of the utility of Corollary 3, consider the channel M in Figure 2, which is defined in [2]. The input alphabet for M is $\{1,2,3,4\}$, and the output alphabet is $\{1,2,3,4,5,6\}$. In section 3 of the supplementary information, we prove that for any 2×2 projection operators $P_1, \ldots, P_6$, the radius of the set $\left\{\sum_{y=1}^6 M(y|x) P_y \mid x = 1,2,3,4\right\}$ is no more than $\frac{1}{6} + \frac{1}{3\sqrt{2}}$. This maximum is achieved when $P_1 = 0$, $P_2 = \mathbb{I}$, and $\{P_3, P_4\}$ and $\{P_5, P_6\}$ are two different Pauli measurements. Therefore,

$$\mathrm{Succ}_{\mathrm{Q}(2)}(M) = \frac{2}{3} + \frac{1}{3\sqrt{2}},$$

and the protocol from [2] is optimal for 2-dimensional entanglement assistance. (We note that this generalizes the paper [8], which showed the optimality of [2] within a more restricted class of protocols.)

	1	2	3	4	5	6
1	1/3	0	1/3	0	1/3	0
2	1/3	0	0	1/3	0	1/3
3	0	1/3	1/3	0	0	1/3
4	0	1/3	0	1/3	1/3	0

FIG. 2. The channel M, from [2].

Non-signaling devices. In order to prove more explicit bounds on the limits of quantum assistance, we first turn our attention to assistance by a non-signaling correlation. The next proposition asserts a formula for the optimal non-signaling assisted success probability of a channel. For any finite set of vectors $S \subseteq \mathbb{R}^k$, let $\mathrm{Rad}_1(S)$ denote the radius of S under the 1-norm.

Proposition 4. *Let N be a classical channel, and for each $x \in \mathcal{X}$, let $n_x = \{N(y \mid x)\}_{y\in\mathcal{Y}} \in \mathbb{R}^{\mathcal{Y}}$. Then,*

$$\mathrm{Succ}_{\mathrm{NS}}(N) = \frac{1}{2} + \frac{1}{2} \cdot \mathrm{Rad}_1\{n_x \mid x \in \mathcal{X}\}. \qquad (4)$$

Note that in the above formula, we take the radius of $\{n_x\}$ *as a subset of* $\mathbb{R}^{\mathcal{Y}}$, not as a subset of the set of probability distributions on $\mathcal{Y}$.

Proof of Proposition 4. By Proposition 14 from [5],

$$\begin{aligned}\mathrm{Succ}_{\mathrm{NS}}(N) &= 1 - \max_{c\in\mathbb{R}^{\mathcal{Y}}} \min_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} (\min\{c_y, N(y \mid x)\} - c_y/2) \\ &= \min_{c\in\mathbb{R}^{\mathcal{Y}}} \max_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} (1 - \min\{c_y, N(y \mid x)\} + c_y/2).\end{aligned}$$

Using the easily proved fact that $\|u - v\|_1 = \sum_i u_i + \sum_i v_i - 2\sum_i \min\{u_i, v_i\}$, the above formula simplifies to

$$\mathrm{Succ}_{\mathrm{NS}}(N) = \min_{c\in\mathbb{R}^{\mathcal{Y}}} \max_{x\in\mathcal{X}} \left(\frac{1}{2} + \frac{1}{2} \cdot \|c - n_x\|_1\right),$$

which implies (4) above. □

Formula (4) allows us to relate the quantity $\mathrm{Succ}_{\mathrm{NS}}(N)$ to the quantity $\mathrm{Succ}(N)$.

Theorem 5. *Let N be a classical channel, and let $r = |\mathcal{X}|$ denote the size of the input alphabet of N. Then,*

$$\mathrm{Succ}_{\mathrm{NS}}(N) - \frac{1}{2} \leq \left(2 - \frac{2}{r}\right)\left[\mathrm{Succ}(N) - \frac{1}{2}\right]. \qquad (5)$$

Proof. Let $\{n_x\}$ be the vectors defined in Proposition 4. The unassisted one-shot success probability can be expressed in terms of these vectors like so:

$$\mathrm{Succ}(N) = \frac{1}{2} + \frac{1}{4} \cdot \mathrm{Diam}_1\{n_x\}, \qquad (6)$$

where Diam_1 denotes diameter under the 1-norm. A triangle-inequality argument shows that the distance from the mean vector $\left(\sum_x n_x\right)/r$ to the set $\{n_x\}$ cannot exceed $\left(1 - \frac{1}{r}\right)\mathrm{Diam}_1\{n_x\}$. Therefore, $\mathrm{Rad}_1\{n_x\} \leq \left(1 - \frac{1}{r}\right)\mathrm{Diam}_1\{n_x\}$, which implies the desired result. □

Theorem 5 is the best possible in the sense that there are channels where equality is achieved in (5). Consider the following example, which is a generalization of the channel M from Figure 2. Let s be a positive integer. For any $i \in \{0,1,2,\ldots,2^s-1\}$, let $b_i \in \mathbb{F}_2^s$ denote the binary representation of i. Define a channel T as follows. The input alphabet of T is $\{0,1,2,\ldots,2^s-1\}$, and the output alphabet is $\{1,2,\ldots,2^s-1\}\times\{0,1\}$. On given input i, the channel chooses an element $j \in \{1,\ldots,2^s-1\}$ uniformly at random and outputs the pair $(j, b_i \cdot b_j)$ (where $b_i \cdot b_j$ denotes the inner product of b_i and b_j mod 2).

For any $i \in \{0,1,2,\ldots,2^s-1\}$, let $\ell_i \in \mathbb{R}^{2(2^s-1)}$ denote the probability vector which expresses the output of T on input i. It is easy to see that the diameter of $\{\ell_i\}$ is $2^s/(2^s-1)$, and thus $\mathrm{Succ}(T) = \frac{1}{2} + 2^{s-2}/(2^s-1)$. On

the other hand, the radius of $\{\ell_i\}$ is 1, as can be seen from the following calculation. For any $c \in \mathbb{R}^{2(2^s-1)}$,

$$\begin{aligned}
&\max_{0\le i\le 2^s-1} \|\ell_i - c\|_1 \ge 2^{-s} \sum_{i=0}^{2^s-1} \|\ell_i - c\|_1 \\
&= 2^{-s} \sum_{\substack{1\le j\le 2^s-1 \\ t\in\{0,1\}}} \left[2^{s-1}\left|c_{jt} - (2^s-1)^{-1}\right| + 2^{s-1}\left|c_{jt} - 0\right|\right] \\
&\ge 2^{-s} \sum_{\substack{1\le j\le 2^s-1 \\ t\in\{0,1\}}} \left[2^{s-1}(2^s-1)^{-1}\right] = 1.
\end{aligned}$$

Thus $\mathrm{Succ}_{\mathrm{NS}}(T) = 1$. (And, indeed, a perfect communication protocol for T exists–see section 4 of the supplementary information.) The channel T achieves equality in (5).

The following modifiied version of Theorem 5 will be useful in our analysis of entanglement assistance.

Theorem 6. *Let N be a classical channel, and let D be a non-signaling correlation arising from a two-part device (D_A, D_B). Let m denote the size of the output alphabet of D_A. Then,*

$$\mathrm{Succ}(N, D) - \frac{1}{2} \le \left(2 - \frac{1}{m}\right)\left[\mathrm{Succ}(N) - \frac{1}{2}\right]. \quad (7)$$

Proof. A protocol for communicating a single bit a using N and D proceeds as follows. Alice uses a to choose an input to D_A, and then uses a and the output of D_A to choose an input to N. Bob uses the output of N to choose an input to D_B, and then uses the outputs of N and D_B together to guess the bit a.

The optimal success probability $\mathrm{Succ}(N, D)$ can be achieved by a *deterministic* protocol (i.e., a protocol in which Alice and Bob make their choices according to deterministic functions). As there are only $2m$ possible inputs that Alice could make to N in a deterministic protocol, the success probability of such a protocol is bounded by $(2 - 2/(2m))\mathrm{Succ}(N)$ by Theorem 5. □

Binary quantum devices. Finally, we will use our bounds for non-signaling devices to obtain bounds for assistance by binary quantum devices.

A two-part device D is *local-deterministic* if the output of each part is a deterministic function of its input. A non-signaling correlation is *local* if it is a convex combination of local-deterministic correlations. We define the *local fraction* of a non-signaling correlation, a concept which is used in [9], [10].

Definition 7. *Let D be a non-signaling correlation. The local fraction of D, denoted $\mathrm{loc}(D)$, is the largest real number $\alpha \in [0, 1]$ such that there exists a decomposition*

$$D = \alpha L + (1 - \alpha)F, \quad (8)$$

where L is a local correlation and F is a non-signaling correlation.

For any classical channel N, it is easy to see that when a decomposition such as (8) exists with L local and F non-signaling,

$$\begin{aligned}
\mathrm{Succ}(N, D) &\le \alpha\mathrm{Succ}(N, L) + (1 - \alpha)\mathrm{Succ}(N, F) \\
&\le \alpha\mathrm{Succ}(N) + (1 - \alpha)\mathrm{Succ}_{\mathrm{NS}}(N).
\end{aligned}$$

This implies the following stronger version of Theorem 6.

Theorem 8. *Let N be a channel, and let D be a non-signaling correlation arising from a two-part device (D_A, D_B). Let m denote the size of the output alphabet of D_A. Then*

$$\frac{\mathrm{Succ}(N, D) - \frac{1}{2}}{\mathrm{Succ}(N) - \frac{1}{2}} \le 1 + \left(1 - \frac{1}{m}\right)(1 - \mathrm{loc}(D)). \quad \square$$

Thus, to obtain improved upper bounds on $\mathrm{Succ}(N, D)$ for quantum correlations D, it suffices to find lower bounds on the local fractions of quantum correlations. In section 5 of the supplementary material, we use facts about the geometry of quantum and non-signaling correlations [11] to prove the following bound for binary quantum correlations.

Proposition 9. *Let D be a binary quantum correlation. Then $\mathrm{loc}(D) \ge 2 - \sqrt{2}$.*

Combining Theorem 8 and Proposition 9 yields the following.

Corollary 10. *For any classical channel N,*

$$\frac{\mathrm{Succ}_{\mathrm{Q}_b}(N) - \frac{1}{2}}{\mathrm{Succ}(N) - \frac{1}{2}} \le \frac{1}{2} + \frac{1}{\sqrt{2}}.$$

Note that equality occurs in Corollary 10 for the case discussed in [2].

Conclusion. We have given a formula for the n-dimensional entanglement-assisted one-shot success probability of a classical channel, and have shown its utility by using it to show that the protocol in [2] is optimal. We derived a more explicit bound on the advantage gained by binary quantum correlations (which is an equality in the case of [2]). Along the way, we established a bound on the advantage gained by non-signaling assistance and provided an example where equality is achieved.

Future research could explore methods for evaluating the formula from Theorem 2. (Section 3 of the supplementary information provides methods which might generalize.) Also, it would be interesting to try to prove stronger bounds on the increase in $\mathrm{Succ}(N)$ that is provided by entanglement. (This might involve generalizations of Proposition 9.) Another natural next step would be to consider the one-shot success probability for non-binary messages.

The authors would like to thank Vincent Russo for his help with the preparation and editing of this paper. We

also thank Shmuel Friedland, Aubrey da Cunha, Xiaodi Wu, and Kim Winick for many useful discussions, and the anonymous PRL reviewers for helpful comments. This research was supported in part by the National Basic Research Program of China under Awards 2011CBA00300 and 2011CBA00301, and the NSF of the United States under Award 1017335.

* bhemen@umich.edu
† carlmi@umich.edu
‡ shiyy@umich.edu
§ wootters@umich.edu

[1] T. Cubitt, D. Leung, W. Matthews, and A. Winter, Physical Review Letters **104**, 230503 (2010).
[2] R. Prevedel, Y. Lu, W. Matthews, R. Kaltenback, and K. J. Resch, Physical Review Letters **106**, 110505 (2011).
[3] T. Cubitt, D. Leung, W. Matthews, and A. Winter, IEEE Transactions on Information Theory **57**, 5509 (2011).
[4] C. Bennett, P. Shor, J. Smolin, and A. Thapliyal, IEEE Transactions on Information Theory **48**, 2637 (2002).
[5] W. Matthews, IEEE Transactions on Information Theory **58**, 7036 (2012).
[6] S. Beigi, Physical Review A **82**, 010303 (2010).
[7] D. Leung, L. Mancinska, W. Matthews, M. Ozols, and A. Roy, Communications in Mathematical Physics **311**, 97 (2012), ISSN 0010-3616, URL http://dx.doi.org/10.1007/s00220-012-1451-x.
[8] H. Williams and P. Bourdon, Arxiv preprint arXiv:1109.1029 (2011).
[9] A. Elitzur, S. Popescu, and D. Rohrlich, Physics Letters A **162**, 25 (1992).
[10] J. Barrett, A. Kent, and S. Pironio, Physical Review Letters **97**, 170409 (2006).
[11] B. S. Tsirel'son, Hadronic Journal Supplement **8**, 329 (1993).

SUPPLEMENTARY INFORMATION

1. The radius of a set of Hermitian operators

Lemma 1. *For any finite set $\{H_i\}_{i\in\mathcal{I}}$ of Hermitian operators on a finite-dimensional Hilbert space V, the radius of $\{H_i\}$ is equal to*

$$\max_{\substack{\lambda_i\geq 0,\lambda_i'\geq 0\\ \sum\lambda_i=\sum\lambda_i'\\ \mathrm{Tr}(\sum\lambda_i)=1/2}} \left[\sum_{i\in\mathcal{I}} \mathrm{Tr}\left((\lambda_i-\lambda_i')H_i\right)\right].$$

Proof. Any family of Hermitian operators $\{H_i\}$ may be translated to a family $\{H_i+W\}$ which contains the operator 0. This translation does not affect the radius nor the expression from the statement of the lemma. Therefore, we may assume that $\{H_i\}$ contains 0. By definition,

$$\mathrm{Rad}\{H_i\}_i \quad = \quad \min_{\substack{C,r\\ C-H_i\geq -r\mathbb{I}\\ H_i-C\geq -r\mathbb{I}}} (r), \tag{1}$$

where the maximization is over Hermitian operators C and real numbers r. Since $0\in\{H_i\}$, whenever the constraints in this maximization are satisfied we have in particular that $C\geq -r\mathbb{I}$. Letting $Z=C+r\mathbb{I}$, we obtain the following alternate expression:

$$\mathrm{Rad}\{H_i\}_i \quad = \quad \min_{\substack{Z,r\\ Z\geq H_i\\ -Z+2r\mathbb{I}\geq -H_i}} (r). \tag{2}$$

By semidefinite programming duality, this is equivalent to

$$\mathrm{Rad}\{H_i\}_i \quad = \quad \max_{\substack{\lambda_i\geq 0,\lambda_i'\geq 0,\\ \sum_i\lambda_i-\sum_i\lambda_i'\leq 0\\ 2\mathrm{Tr}(\sum\lambda_i')\leq 1}} \left[\left(\sum \mathrm{Tr}(\lambda_i H_i)-\sum \mathrm{Tr}(\lambda_i' H_i)\right)\right].$$

It is easy to see that this maximum is achieved by a pair of families $\{\lambda_i\},\{\lambda_i'\}$ satisfying $\sum\lambda_i=\sum\lambda_i'$ and $2\mathrm{Tr}(\sum\lambda_i')=1$. □

2. The proof of Corollary 3 in the main text

The radius function is convex in the following sense: for any familes of operators $\{J_y\}_{y\in\mathcal{Y}}$ and $\{K_y\}_{y\in\mathcal{Y}}$, and real number $\alpha\in[0,1]$,

$$\begin{aligned}&\mathrm{Rad}\{\alpha J_y+(1-\alpha)K_y\}_y\\ &\qquad\leq \alpha\mathrm{Rad}\{J_y\}_y+(1-\alpha)\mathrm{Rad}\{K_y\}_y.\end{aligned}$$

(For, if we let J' be such that the distance from J' to $\{J_y\}_y$ is equal to $r:=\mathrm{Rad}\{J_y\}_y$, and we let K' be such that the distance from K' to $\{K_y\}_y$ is equal to $r':=\mathrm{Rad}\{K_y\}_y$, then the distance from $\alpha J'+(1-\alpha)K'$ to $\{\alpha J_y+(1-\alpha)K_y\}_y$ is no more than $\alpha r+(1-\alpha)r'$ by the triangle inequality.) In particular, this

convexity property implies that the radius of $\{\alpha J_y + (1-\alpha)K_y\}_y$ is no more than the maximum of $\mathrm{Rad}\{J_y\}_y$ and $\mathrm{Rad}\{K_y\}_y$.

Since any Hermitian operator B satisfying $0 \leq B \leq \mathbb{I}$ is a convex combination of projection operators, Corollary 3 follows from Theorem 2.

3. An example calculation

Let M be the channel defined in figure 2 in the main text. In this section we will use Theorem 2 from the main text to calculate the quantity $\mathrm{Succ}_{\mathrm{Q}(2)}(M)$. First, we will prove the following lemma which provides a simplified formula for $\mathrm{Succ}_{\mathrm{Q}(n)}(M)$. For any projection operator P, let $P^\perp$ denote projection onto the orthogonal complement of P.

Lemma 2. *For any $n \geq 1$, the quantity $\mathrm{Succ}_{\mathrm{Q}(n)}(M)$ is equal to*

$$\frac{1}{2} + \left(\frac{1}{3}\right) \max_{X,Y,Z} \left(\mathrm{Rad}\left\{X+Y+Z, X+Y^\perp+Z^\perp, X^\perp+Y+Z^\perp, X^\perp+Y^\perp+Z\right\}\right),$$

where the maximum is taken over all projection operators X, Y, Z on $\mathbb{C}^n$.

Proof. For any Hermitian operators $B_1, B_2, B_3, B_4, B_5, B_6$ on $\mathbb{C}^n$, let

$$F(B_1, B_2, B_3, B_4, B_5, B_6)$$

be equal to the quantity

$$\mathrm{Rad}\left\{B_1+B_3+B_5, B_1+B_4+B_6, B_2+B_3+B_6, B_2+B_4+B_5\right\}.$$

By the formula from Theorem 2 in the main text,

$$\mathrm{Succ}_{\mathrm{Q}(n)}(M) = \frac{1}{2} + \left(\frac{1}{3}\right) \max_{0 \leq B_i \leq \mathbb{I}} F(B_1, B_2, B_3, B_4, B_5, B_6). \tag{3}$$

Let

$$m = \max_{0 \leq B_i \leq \mathbb{I}} F(B_1, B_2, B_3, B_4, B_5, B_6). \tag{4}$$

It suffices to prove that this maximum is achieved by some 6-tuple of the form $(X, X^\perp, Y, Y^\perp, Z, Z^\perp)$, where X, Y, and Z are projections.

As noted in section 2 of the supplementary information, the radius function is convex in the sense that if (H_1, H_2, H_3, H_4) and (H_1', H_2', H_3', H_4') are Hermitian operators and $\alpha \in [0,1]$ is a real number,

$$\mathrm{Rad}\{\alpha H_i + (1-\alpha)H_i'\}_i \leq \alpha \mathrm{Rad}\{H_i\}_i + (1-\alpha)\mathrm{Rad}\{H_i'\}_i. \tag{5}$$

It follows easily by linearity that a similar convexity property holds for F: for any Hermitian operators $B_1, \ldots, B_6$ and $B_1', \ldots, B_6'$, and any $\alpha \in [0,1]$,

$$\begin{aligned} & F(\alpha B_1 + (1-\alpha)B_1', \ldots, \alpha B_6 + (1-\alpha)B_6') \\ \leq\ & \alpha F(B_1, \ldots, B_6) + (1-\alpha)F(B_1', \ldots, B_6'). \end{aligned}$$

In particular,

$$\begin{aligned} & F(\alpha B_1 + (1-\alpha)B_1', \ldots, \alpha B_6 + (1-\alpha)B_6') \\ \leq\ & \max\left\{F(B_1, \ldots, B_6), F(B_1', \ldots, B_6')\right\}. \end{aligned} \tag{6}$$

Additionally, F is translation-invariant in the following sense: for any Hermitian operators $B_1, \ldots, B_6$, and any Hermitian operators K, L, and M,

$$F(B_1+K, B_2+K, B_3+L, B_4+L, B_5+M, B_6+M) = F(B_1, \ldots, B_6). \tag{7}$$

Let $X_1, X_2, Y_1, Y_2, Z_1, Z_2$ be Hermitian operators satisfying $0 \leq X_i, Y_i, Z_i \leq \mathbb{I}$ such that $F(X_1, X_2, Y_1, Y_2, Z_1, Z_2) = m$. Let X_+ and X_- be a pair of positive semidefinite operators having mutual orthogonal supports which are such that

$$X_1 - X_2 = X_+ - X_-. \tag{8}$$

Define Y_+, Y_-, Z_+, Z_- similarly. By property (7) above,

$$\begin{aligned} F(X_+, X_-, Y_+, Y_-, Z_+, Z_-) &= F(X_1, X_2, Y_1, Y_2, Z_1, Z_2) \\ &= m. \end{aligned} \tag{9}$$

The pair (X_+, X_-) can be expressed as a convex combination of pairs of projections $(P_1^{(i)}, P_2^{(i)})$ where for each i, the support of $P_1^{(i)}$ is orthogonal to $P_2^{(i)}$. A similar decomposition exists for (Y_+, Y_-) and (Z_+, Z_-). Therefore by property (6) above, there exist pairs of projections (P_1, P_2), (Q_1, Q_2), (R_1, R_2), with each pair having mutually orthogonal supports, such that

$$F(P_1, P_2, Q_1, Q_2, R_1, R_2) = m. \tag{10}$$

Let $P_3 = \mathbb{I} - P_1 - P_2$, and define Q_3 and R_3 similarly. By (7),

$$F\left(P_1 + \frac{P_3}{2}, P_2 + \frac{P_3}{2}, Q_1 + \frac{Q_3}{2}, Q_2 + \frac{Q_3}{2}, R_1 + \frac{R_3}{2}, R_2 + \frac{R_3}{2}\right) = m. \tag{11}$$

The 6-tuple on the left hand side of the equation above is a convex combination of the 6-tuples

$$\begin{gathered} (P_1 + P_3, P_2, Q_1 + Q_3, Q_2, R_1 + R_3, R_2) \\ \text{and } (P_1, P_2 + P_3, Q_1, Q_2 + Q_3, R_1, R_2 + R_3)\,. \end{gathered}$$

By (6), at least one of these 6-tuples must achieve the maximum m. This completes the proof. □

Lemma 3. *For any projection operators X, Y, Z on the two-dimensional vector space $\mathbb{C}^2$, the radius of the set*

$$\left\{X + Y + Z, X + Y^\perp + Z^\perp, X^\perp + Y + Z^\perp, X^\perp + Y^\perp + Z\right\}. \tag{12}$$

is less than or equal to $\frac{1}{2} + \frac{1}{\sqrt{2}}$.

Proof. **Case 1: The matrices X, Y, and Z are all scalar matrices.** In this case, each of X, Y, and Z is equal to either 0 or $\mathbb{I}$. This case is trivial, since the radius of the set $\{3\mathbb{I}, \mathbb{I}\}$ is 1, and the radius of the set $\{2\mathbb{I}, 0\}$ is 1.

Case 2: Two of the matrices X, Y, Z are scalar matrices and one is a nonscalar. We may assume without loss of generality that X is the nonscalar matrix. Then the set (12) is equal to either

$$\{0, X + \mathbb{I}, 2\mathbb{I}\} \tag{13}$$

or

$$\{X, X + 2\mathbb{I}, \mathbb{I}\}\,. \tag{14}$$

In the former case, the operator-norm distance from the operator $\mathbb{I}$ to the set $\{0, X + \mathbb{I}, 2\mathbb{I}\}$ is 1. In the latter case, the operator-norm distance from the operator $X + \mathbb{I}$ to the set $\{X, X + 2\mathbb{I}, \mathbb{I}\}$ is 1. The desired result follows.

Case 3: Exactly one of the matrices X, Y, Z is a scalar matrix. We may assume that X and Y are nonscalar matrices and Z is scalar. Also, by replacing (X, Y, Z) with $(X^\perp, Y, Z^\perp)$ if necessary, we may assume that $Z = \mathbb{I}$.

Let $X = |x\rangle\langle x|$ and $Y = |y\rangle\langle y|$ where $x, y \in \mathbb{C}^2$ are unit vectors, and let $\theta = \arccos(|x \cdot y|)$. Both of the operators

$$X + Y + \mathbb{I}, X^\perp + Y^\perp + \mathbb{I} \tag{15}$$

have eigenvalues $\{2 + \cos\theta, 2 - \cos\theta\}$, and both of the operators

$$X + Y^\perp, X^\perp + Y \tag{16}$$

have eigenvalues $\{1 + \sin\theta, 1 - \sin\theta\}$. If we let

$$C = \left(\frac{3}{2} + \frac{\cos\theta - \sin\theta}{2}\right)\mathbb{I}, \tag{17}$$

then the operator norm distance from C to each of the elements of (12) is $\frac{1}{2} + \frac{\cos\theta+\sin\theta}{2} \leq \frac{1}{2} + \frac{1}{\sqrt{2}}$.

Case 4: Each of X, Y, Z is a nonscalar matrix. As in case 3, let $X = |x\rangle\langle x|$ and $Y = |y\rangle\langle y|$ and let $\theta = \arccos(|x \cdot y|)$.

Let

$$C = \mathbb{I} + \left(\frac{1}{2} + \frac{\cos\theta - \sin\theta}{2}\right)Z + \left(\frac{1}{2} - \frac{\cos\theta - \sin\theta}{2}\right)Z^\perp. \tag{18}$$

Then, the operator norm of the difference

$$(X + Y + Z) - C = (X + Y) - \left(\frac{3}{2} + \frac{\cos\theta - \sin\theta}{2}\right)\mathbb{I} \tag{19}$$

is $\frac{1}{2} + \frac{\cos\theta+\sin\theta}{2}$, which is less than or equal to $\frac{1}{2} + \frac{1}{\sqrt{2}}$. A similar calculation shows that the distance from C to each of the other three elements of set (12) is equal to $\frac{1}{2} + \frac{\cos\theta+\sin\theta}{2}$. This completes the proof. □

For any angle $\theta \in \mathbb{R}$, let $P_\theta \colon \mathbb{C}^2 \to \mathbb{C}^2$ denote projection onto the unit vector $\cos(\theta)|0\rangle + \sin(\theta)|1\rangle$. Consider the set

$$\left\{P_0 + P_{\pi/4} + \mathbb{I}, P_0 + P_{3\pi/4}, P_{\pi/2} + P_{\pi/4}, P_{\pi/2} + P_{3\pi/4} + \mathbb{I}\right\} \tag{20}$$

A direct calculation shows that the distance from the operator $\left(\frac{3}{2}\right)\mathbb{I}$ to set (20) is $\frac{1}{2} + \frac{1}{\sqrt{2}}$. The next lemma asserts that this quantity is in fact the radius of (20).

Lemma 4. *The radius of the set*

$$\left\{P_0 + P_{\pi/4} + \mathbb{I}, P_0 + P_{3\pi/4}, P_{\pi/2} + P_{\pi/4}, P_{\pi/2} + P_{3\pi/4} + \mathbb{I}\right\} \tag{21}$$

is is equal to $\frac{1}{2} + \frac{1}{\sqrt{2}}$.

Proof. For any Hermitian operator $H \colon \mathbb{C}^2 \to \mathbb{C}^2$, let us write $\overline{H}$ to denote the trace-zero operator $H - (\mathrm{Tr}(H))\mathbb{I}/2$. In the proof that follows, we will make use of the following fact: for any two Hermitian operators $Q, R \colon \mathbb{C}^2 \to \mathbb{C}^2$,

$$||Q - R|| = |\mathrm{Tr}(Q) - \mathrm{Tr}(R)| + \left|\left|\overline{Q} - \overline{R}\right|\right| \tag{22}$$

Suppose, for the sake of contradiction, that there exists a Hermitian operator Z whose distance from each of the elements of set (20) is strictly less than $\frac{1}{2} + \frac{1}{\sqrt{2}}$. Then,

$$\begin{aligned} 2\left(\frac{1}{2}+\frac{1}{\sqrt{2}}\right) &> \left\|(P_0 + P_{3\pi/4}) - Z\right\| + \left\|(P_{\pi/2} + P_{\pi/4}) - Z\right\| \\ &= \left\|\left(P_0 + P_{3\pi/4} - \mathbb{I}\right) - \overline{Z}\right\| + \left\|\left(P_{\pi/2} + P_{\pi/4} - \mathbb{I}\right) - \overline{Z}\right\| + 2\cdot|2 - \mathrm{Tr}(Z)| \\ &\geq \left\|\left(P_0 + P_{3\pi/4}\right) - \left(P_{\pi/2} + P_{\pi/4}\right)\right\| + 2\cdot|2 - \mathrm{Tr}(Z)| \\ &= \sqrt{2} + 2\cdot|2 - \mathrm{Tr}(Z)| \end{aligned}$$

Therefore, $\mathrm{Tr}(Z) < \frac{5}{2}$. Similarly,

$$\begin{aligned} 2\left(\frac{1}{2}+\frac{1}{\sqrt{2}}\right) &> \left\|(P_0 + P_{\pi/4} + \mathbb{I}) - Z\right\| + \left\|(P_{\pi/2} + P_{3\pi/4}) - \mathbb{I}\right\| \\ &= \left\|\left(P_0 + P_{\pi/4} - \mathbb{I}\right) - \overline{Z}\right\| + \left\|\left(P_{\pi/2} + P_{3\pi/4} - \mathbb{I}\right) - \overline{Z}\right\| + 2\cdot|3 - \mathrm{Tr}(Z)| \\ &\geq \left\|\left(P_0 + P_{\pi/4}\right) - \left(P_{\pi/2} + P_{3\pi/4}\right)\right\| + 2\cdot|3 - \mathrm{Tr}(Z)| \\ &= \sqrt{2} + 2\cdot|3 - \mathrm{Tr}(Z)|\,, \end{aligned}$$

which implies $\mathrm{Tr}(Z) > \frac{5}{2}$. This is a contradiction. □

Combining Lemmas 2–4, we have the following proposition.

Proposition 5. *The quantity* $\mathrm{Succ}_{\mathrm{Q}(2)}(M)$ *is equal to* $\frac{2}{3} + \frac{1}{3\sqrt{2}}$. □

4. An example of optimal non-signaling assistance

In this section we discuss an example in which equality occurs in Theorem 5 from the main text. This example is a generalization of the protocol from [2].

Let m be a positive integer. Let

$$\mathcal{Z} = \mathbb{F}_2^m, \tag{23}$$

$$\mathcal{W} = (\mathbb{F}_2^m \setminus \{0\}) \times \mathbb{F}_2. \tag{24}$$

Let K be a channel defined as follows:

(1) The input alphabet of K is $\mathcal{Z}$, and the output alphabet of K is $\mathcal{W}$.
(2) For any given input $\mathbf{v} \in \mathbb{F}_2^m$, the output of K is unformly distributed over the set

$$\{(\mathbf{w}, \mathbf{w}\cdot\mathbf{v}) \mid \mathbf{w} \in \mathbb{F}_2^m \setminus \{0\}\}. \tag{25}$$

(Here, $\mathbf{w}\cdot\mathbf{v} \in \mathbb{F}_2$ denotes the inner product of $\mathbf{w}$ and $\mathbf{v}$.)

Let (E_1, E_2) be a two part input-output device defined as follows. (See Figure 1.)

(1) The input alphabet for E_1 is $\mathbb{F}_2$, and the output alphabet for E_1 is $\mathcal{Z}$.
(2) The input alphabet for E_2 is $\mathcal{W}$, and the output alphabet for E_2 is $\mathbb{F}_2$.
(3) If the inputs to E_1 and E_2 are $a \in \{0,1\}$ and $(\mathbf{w}, r) \in (\mathbb{F}_2^m \setminus \{\mathbf{0}\}) \times \mathbb{F}_2$, then the output of E_1 is unformly distributed over all vectors $\mathbf{a} = (a_1, a_2, \ldots, a_m)$ that satisfy $a_1 = a$, and the output of E_2 is $a \oplus r \oplus (\mathbf{w}\cdot\mathbf{a})$.

It can be checked that the correlation E arising from (E_1, E_2) is non-signaling. Additionally, one can see (by substitution) that using E to assist K yields a perfect transmission of a single bit. (See figure 2.)

Now, let us calculate the quantity $\mathrm{Succ}\,(K)$. For any two distinct vectors $\mathbf{x}_0, \mathbf{x}_1 \in \mathbb{F}_2^m$, the probability that a randomly chosen vector $\mathbf{w} \in \mathbb{F}_2^m \setminus \{\mathbf{0}\}$ will satisfy

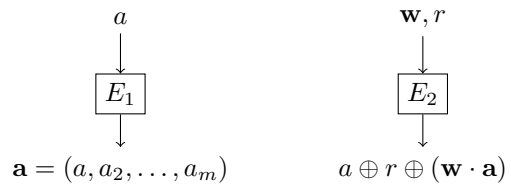

FIGURE 1. The device (E_1, E_2).

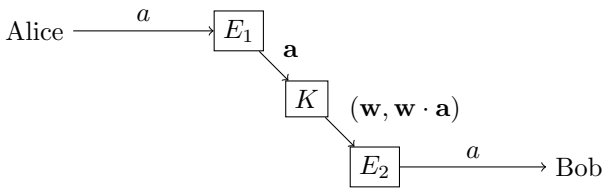

FIGURE 2. A perfect communication protocol.

$\mathbf{w}\cdot\mathbf{x}_0 \neq \mathbf{w}\cdot\mathbf{x_1}$ is equal to $2^{m-1}/\left(2^m-1\right)$. This fact has the following consequence: if Alice employs the deterministic encoding strategy $[0 \mapsto \mathbf{x}_0, 1 \mapsto \mathbf{x}_1]$ to send a single bit, then the optimal probability with which Bob can decode is

$$\left[\frac{2^{m-1}}{2^m-1}\right](1) + \left[\frac{2^{m-1}-1}{2^m-1}\right]\left(\frac{1}{2}\right) \tag{26}$$

$$= \quad \frac{2^m + 2^{m-1} - 1}{2^{m+1}-2}. \tag{27}$$

Therefore, $\mathrm{Succ}\left(K\right)$ is equal to quantity (27), while $\mathrm{Succ}_{\mathrm{NS}}\left(K\right)$ is equal to 1. Theorem 5 from the main text asserts the following bound on $\mathrm{Succ}_{\mathrm{NS}}\left(K\right)$:

$$\begin{aligned}
\mathrm{Succ}_{\mathrm{NS}}\left(K\right) &\leq \frac{1}{2} + \left(2 - \frac{2}{2^m}\right)\left[\mathrm{Succ}(K) - \frac{1}{2}\right] \\
&= \frac{1}{2} + 2\left(\frac{2^m-1}{2^m}\right)\left(\frac{2^{m-1}}{2^{m+1}-2}\right) \\
&= 1.
\end{aligned}$$

Therefore, equality is achieved in Theorem 5 from the main text when $N = K$.

5. The local fraction of a binary quantum correlation

In this section, we prove the following proposition from the main text.

Proposition 6. *Let D be a binary quantum correlation. Then* $\mathrm{loc}(D) \geq 2 - \sqrt{2}$.

Proof. For any binary non-signaling correlation G, let

$$f_1\left(G\right) = \sum_{a,x,b,y\in\{0,1\}} (-1)^{x\oplus b\oplus(a\wedge y)} G(xb|ay).$$

This is the function which defines the CHSH inequality [1]. let f_2, f_3, and f_4 be the functions defined by the same expression with $a \wedge y$ replaced by $\neg a \wedge y$, $a \wedge \neg y$, and $\neg a \wedge \neg y$, respectively.

We note the following facts. (See [3].)

(1) A non-signaling correlation G is local if and only if $-2 \leq f_i(G) \leq 2$ for $i = 1, 2, 3, 4$.

(2) If G is a quantum correlation, then for $i = 1, 2, 3, 4$,
$$-2\sqrt{2} \leq f_i(G) \leq 2\sqrt{2}.$$

(3) There are eight non-signaling correlations $\{P_i^+\}_{i=1}^4$ and $\{P_i^-\}_{i=1}^4$, satisfying
$$f_j\left(P_i^{\pm}\right) = \begin{cases} \pm 4 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}$$
These are the *Popescu-Rohrlich* (PR) boxes.

(4) Every non-signaling correlation is a convex combination of local correlations and the eight PR boxes. Further, for any two distinct PR boxes P and P', the correlation $(P + P')/2$ is local.

From the second part of item 4, it follows that any convex combination of local boxes and PR boxes can be simplified into an expression of the form $\alpha L + (1-\alpha)Q$, where L is local, Q is a PR box, and $\alpha \in [0, 1]$. Any non-signaling correlation can thus be expressed as a convex combination of a local correlation and a single PR box.

Let $D = \alpha L + (1 - \alpha)Q$, where L is local and Q is a PR box. First suppose that $Q = P_j^+$. Let $L_\beta = (\alpha L + (\beta - \alpha)P_j^+)/\beta$. for any $\beta \in [\alpha, 1]$. Then L_β is local whenever $f_j(L_\beta) \leq 2$. If $f_j(L_1) < 2$, then $L_1 (= D)$ is local, and the proposition follows easily. Otherwise, there is a value $\beta \in [\alpha, 1]$ such that $f_j(L_\beta) = 2$. We have $D = \beta \cdot L_\beta + (1 - \beta)P_j^+$. The quantity β must be at least $2 - \sqrt{2}$, since otherwise (2) would be violated. Therefore $\mathrm{loc}(D) \geq 2 - \sqrt{2}$.

A similar argument completes the proof in the case where $Q = P_j^-$. □

References

[1] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880–884, 1969.

[2] R. Prevedel, Y. Lu, W. Matthews, R. Kaltenback, and K. J. Resch. Entanglement-enhanced classical communication over a noisy classical channel. *Physical Review Letters*, 106:110505, 2011.

[3] B. S. Tsirel'son. Some results and problems on quantum bell-type inequalities. *Hadronic Journal Supplement*, 8(4):329–345, 1993.